

A comparison of bias approximations for the 2SLS estimator

Maurice J. G. Bun
Frank Windmeijer

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP07/10

A Comparison of Bias Approximations for the 2SLS Estimator

MAURICE J.G. BUN* AND FRANK WINDMEIJER[†]

April 2010

Abstract

We consider the bias of the 2SLS estimator in the linear instrumental variables regression with one endogenous regressor only. By using asymptotic expansion techniques we approximate 2SLS coefficient estimation bias under various scenarios regarding the number and strength of instruments. The resulting approximation encompasses existing bias approximations, which are valid in particular cases only. Simulations show that the developed approximation gives an accurate description of the 2SLS bias in case of either weak or many instruments or both.

JEL codes: C13, C31

Keywords: instrumental variables, 2SLS bias, weak instruments

*Tinbergen Institute and Amsterdam School of Economics, University of Amsterdam, The Netherlands, email: m.j.g.bun@uva.nl.

[†]Department of Economics, University of Bristol, and CEMMAP/IFS, United Kingdom, email: f.windmeijer@bristol.ac.uk.

1. Introduction

It is well known that the two-stage least squares (2SLS) estimator can perform poorly in finite samples when there are weak or many instruments or both. One aspect of this poor accuracy is finite sample bias. When instruments are weak, i.e. only weakly correlated with the endogenous regressors, the 2SLS estimator is biased in the direction of the ordinary least squares (OLS) estimator, see e.g. Bound et al. (1995), Staiger and Stock (1997) and Stock et al. (2002). With many instruments a similar result occurs, see e.g. Bekker (1994).

The approximate bias of the 2SLS estimator can be obtained using ‘higher order’ asymptotics, see Nagar (1959) and Buse (1992). The resulting bias approximation has been derived assuming a finite number of strong instruments. Hahn and Kuersteiner (2002) show that in case of nearly weak instruments this standard higher order approximation is still valid. However, under weak instruments as defined by Staiger and Stock (1997) this approximation breaks down. Cruz and Moreira (2005) show by simulation that in case of weak instruments the standard higher order result is a poor approximation to the true finite sample bias.

An alternative and particular simple bias approximation has been proposed by Hahn and Hausman (2002). They also compare 2SLS bias with that of OLS. There are interesting differences between the standard higher-order 2SLS bias approximation and the alternative Hahn-Hausman approximation. When identification becomes weak, the Hahn-Hausman approximate relative bias (compared with OLS) goes to 1, whereas the higher-order Nagar approximation goes to infinity.

In this paper we analyze the accuracy of both approximations using an as-

ymptotic expansion of the 2SLS estimation error. Following the terminology of Andrews and Stock (2006) we will consider different types of asymptotics, i.e. strong, many, weak and many weak IV asymptotics. We show that the Hahn and Hausman (2002) approximation to the finite sample bias is actually not a higher-order approximation when applying usual strong IV asymptotics. However, it can be considered as a first-order bias approximation (inconsistency) in both the many and many weak IV asymptotics set up. We derive an encompassing higher-order bias approximation of the IV bias for these cases, which in a Monte Carlo study is shown to be more accurate than existing bias approximations.

2. Absolute and Relative 2SLS Bias

As in Hahn and Hausman (2002) and Hahn and Kuersteiner (2002) we consider a simple model specification with one endogenous explanatory variable x_i and k instruments z_i

$$y_i = x_i\beta + u_i \tag{2.1}$$

$$x_i = z_i'\pi + v_i, \tag{2.2}$$

$i = 1, \dots, n$, with

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} \sim IIN \left(0, \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix} \right). \tag{2.3}$$

The 2SLS estimator of β is given by

$$\hat{\beta}_{2SLS} = \frac{x'Z(Z'Z)^{-1}Z'y}{x'Z(Z'Z)^{-1}Z'x} = \beta + \frac{x'Z(Z'Z)^{-1}Z'u}{x'Z(Z'Z)^{-1}Z'x}$$

where y , x and u are the n -vectors $(y_1, \dots, y_n)'$, $(x_1, \dots, x_n)'$ and $(u_1, \dots, u_n)'$ respectively, and Z is the $n \times k$ matrix $(z_1, \dots, z_n)'$.

Using the reduced form specification (2.2) we can write

$$\widehat{\beta}_{2SLS} - \beta = \frac{\pi' Z' u + v' P_Z u}{\pi' Z' Z \pi + 2\pi' Z' v + v' P_Z v}, \quad (2.4)$$

where $P_Z = Z(Z'Z)^{-1}Z'$. For ease of exposition, and following Hahn and Kuersteiner (2002), we assume throughout that z_i are non-stochastic instruments with $\lim_{n \rightarrow \infty} \frac{1}{n} Z'Z$ finite and nonsingular. The approximate bias of the 2SLS estimator can be obtained using ‘higher order’ asymptotics, see Nagar (1959), Buse (1992) and Hahn and Kuersteiner (2002), and is then given by

$$E \left[\widehat{\beta}_{2SLS} \right] - \beta \approx \frac{\sigma_{uv}}{\sigma_v^2} \frac{(k-2)}{\mu}, \quad (2.5)$$

where μ is the concentration parameter,

$$\mu = \frac{\pi' Z' Z \pi}{\sigma_v^2}. \quad (2.6)$$

Hahn and Hausman (2002) derive an alternative approximation to the bias of the 2SLS estimator:

$$\begin{aligned} E \left[\widehat{\beta}_{2SLS} \right] - \beta &\approx \frac{E [\pi' Z' u + v' P_Z u]}{E [\pi' Z' Z \pi + 2\pi' Z' v + v' P_Z v]} \\ &= \frac{\sigma_{uv}}{\sigma_v^2} \frac{k}{\mu + k}. \end{aligned} \quad (2.7)$$

There are interesting differences between the two approximations, especially when we relate 2SLS bias to OLS bias. This relative bias is used by Bound et al. (1995), Staiger and Stock (1997) and Stock and Yogo (2005) to assess whether instruments are weak. The bias of the OLS estimator can be approximated by (see also Hahn and Hausman, 2002)

$$E \left[\widehat{\beta}_{OLS} \right] - \beta \approx \frac{\sigma_{uv}}{\sigma_v^2} \frac{1}{\frac{k}{n} + 1} = \text{plim} \left(\widehat{\beta}_{OLS} \right) - \beta.$$

Then the relative bias based on the Nagar 2SLS bias approximation is approximately given by

$$\frac{E \left[\widehat{\beta}_{2SLS} \right] - \beta}{E \left[\widehat{\beta}_{OLS} \right] - \beta} \approx \frac{(k-2) \left(\frac{\mu}{n} + 1 \right)}{\mu}.$$

Using the Hahn-Hausman bias approximation the relative bias would be approximated as

$$\frac{E \left[\widehat{\beta}_{2SLS} \right] - \beta}{E \left[\widehat{\beta}_{OLS} \right] - \beta} = \frac{k \left(\frac{\mu}{n} + 1 \right)}{\mu + k}.$$

When μ is close to 0, the Hahn-Hausman approximate relative bias goes to 1, whereas the Nagar approximation goes to infinity.

3. Bias Approximation

From (2.4), we have for the 2SLS estimation error

$$\widehat{\beta}_{2SLS} - \beta = \frac{\pi' Z' u + v' P_Z u}{\pi' Z' Z \pi + 2\pi' Z' v + v' P_Z v} \equiv \frac{c}{d}, \quad (3.1)$$

with

$$\begin{aligned} E[c] &= E[\pi' Z' u + v' P_Z u] \\ &= \sigma_{uv} k, \end{aligned}$$

and

$$\begin{aligned} E[d] &= E[\pi' Z' Z \pi + 2\pi' Z' v + v' P_Z v] \\ &= \sigma_v^2 (\mu + k). \end{aligned}$$

We will use the shorthand notation $E[c] = \bar{c}$ and $E[d] = \bar{d}$ to denote these expectations.

Below we will develop a stochastic expansion of the estimation error (3.1) in terms of decreasing order of magnitude in $s = \max(\mu, k)$. Doing so we do not restrict ourselves a priori regarding the number and strengths of instruments used. The conventional expansion (Nagar, 1959; Rothenberg, 1984) assumes $\mu = O(n)$ and $k = O(1)$, hence is only valid when $\mu \rightarrow \infty$ as $n \rightarrow \infty$. Here we are more flexible because we allow either μ or k (or both) to grow with the sample size. The only assumption we make is that $s \rightarrow \infty$ as $n \rightarrow \infty$ thereby ruling out the case of a finite number of weak instruments in the sense of Staiger and Stock (1997).

It can be shown that $Var[c] = O(s)$ and $Var[d] = O(s)$, hence $c = O_p(s^{1/2})$ and $d = O_p(s^{1/2})$. It should be noted that these results hold regardless of the number or strength of instruments used in estimation because the usual O and O_p notation exploited here only denotes the largest order of magnitude. Hence, we allow the moments of c and d to contain components μ and k of different order of magnitude. Once again, the only important assumption is that $s \rightarrow \infty$ as $n \rightarrow \infty$.

We may express the denominator of the estimation error (3.1) as follows:

$$\begin{aligned} d &= \bar{d} + d - \bar{d} \\ &= \bar{d}(1 + \bar{d}^{-1}(d - \bar{d})), \end{aligned}$$

hence we have

$$d^{-1} = \bar{d}^{-1}(1 + \bar{d}^{-1}(d - \bar{d}))^{-1}.$$

Noting that $\bar{d}^{-1}(d - \bar{d}) = O_p(s^{-1/2})$ the second factor expands as follows

$$(1 + \bar{d}^{-1}(d - \bar{d}))^{-1} = 1 - \bar{d}^{-1}(d - \bar{d}) + \bar{d}^{-2}(d - \bar{d})^2 + O_p\left(s^{-\frac{3}{2}}\right).$$

Denoting the numerator of (3.1) as

$$c = \bar{c} + c - \bar{c},$$

and exploiting the expansion of the denominator we write the estimation error as

$$\begin{aligned}\widehat{\beta}_{2SLS} - \beta &= \frac{\bar{c}}{\bar{d}} + \frac{c - \bar{c}}{\bar{d}} - \frac{\bar{c}(d - \bar{d})}{\bar{d}^2} - \frac{(c - \bar{c})(d - \bar{d})}{\bar{d}^2} \\ &\quad + \frac{\bar{c}(d - \bar{d})^2}{\bar{d}^3} + O_p\left(s^{-\frac{3}{2}}\right).\end{aligned}$$

Taking expectations we get

$$E\left[\widehat{\beta}_{2SLS} - \beta\right] = \frac{\bar{c}}{\bar{d}} - \frac{E[(c - \bar{c})(d - \bar{d})]}{\bar{d}^2} + \frac{\bar{c}E[(d - \bar{d})^2]}{\bar{d}^3} + o(s^{-1}). \quad (3.2)$$

Evaluating the remaining expectations on the right hand side we find

$$\begin{aligned}E[(c - \bar{c})(d - \bar{d})] &= E[(\pi'Z'u + v'P_Zu - \sigma_{uw}k)(2\pi'Z'v + v'P_Zv - k\sigma_v^2)] \\ &= 2E[\pi'Z'uv'Z\pi] + \frac{\sigma_{uw}}{\sigma_v^2}E[(v'P_Zv - \sigma_{uw}k)(v'P_Zv - k\sigma_v^2)] \\ &= 2\sigma_{uw}\pi'Z'Z\pi + \frac{\sigma_{uw}}{\sigma_v^2}(\sigma_v^4(k^2 + 2k) - (\sigma_{uw}k + k\sigma_v^2)\sigma_v^2k + k^2\sigma_{uw}\sigma_v^2) \\ &= 2\sigma_{uw}\sigma_v^2(\mu + k),\end{aligned}$$

$$\begin{aligned}E[(d - \bar{d})^2] &= E[(2\pi'Z'v + v'P_Zv - k\sigma_v^2)^2] \\ &= E[(2\pi'Z'v)^2] + E[(v'P_Zv - k\sigma_v^2)^2] \\ &= 4\sigma_v^2\pi'Z'Z\pi + \sigma_v^4(k^2 + 2k) - k^2\sigma_v^4 \\ &= 2\sigma_v^4(2\mu + k),\end{aligned}$$

where we used the normality assumption (2.3) and, hence, that terms involving odd moments have expectation zero. Hence, we find the following bias approximation:

$$E\left[\widehat{\beta}_{2SLS} - \beta\right] = \frac{\sigma_{uw}}{\sigma_v^2} \left(\frac{k}{(\mu + k)} - \frac{2\mu^2}{(\mu + k)^3} \right) + o(s^{-1}). \quad (3.3)$$

4. Comparison with Existing Bias Approximations

The standard approach is to assume a fixed number of strong instruments, i.e. $k = O(1)$ and $\mu = O(n)$. In this case $s = O(n)$, the leading term in the bias approximation (3.3) is $O(n^{-1})$ and 2SLS is a consistent estimator. Because k is of smaller order than μ we have

$$\begin{aligned} \frac{1}{\mu + k} &= \frac{1}{\mu} - \frac{k}{\mu(\mu + k)} \\ &= \frac{1}{\mu} + O(n^{-2}), \end{aligned}$$

$$\begin{aligned} \frac{1}{(\mu + k)^3} &= \frac{1}{\mu^3} - \frac{3\mu k^2 + 3\mu^2 k + k^3}{\mu^3(\mu + k)^3} \\ &= \frac{1}{\mu^3} + O(n^{-4}), \end{aligned}$$

hence we have for the approximate bias

$$E \left[\widehat{\beta}_{2SLS} - \beta \right] = \frac{\sigma_{uw}}{\sigma_v^2} \frac{(k-2)}{\mu} + o(n^{-1}),$$

which is equal to the result of Nagar (1959), Buse (1992) and Hahn and Kuersteiner (2002). From the derivations above it is also clear that the Hahn and Hausman (2002) bias approximation is not a higher order approximation because it omits important $O(n^{-1})$ contributions, while adding some $o(n^{-1})$ contributions.

The next question is in which set up the Hahn and Kuersteiner (2002) is actually a valid bias approximation. One possibility is the case of many instruments (Bekker, 1994). More in particular, assuming that $k \rightarrow \infty$, $n \rightarrow \infty$ and $k/n \rightarrow \alpha$ with $0 < \alpha < 1$ we have that $k = O(n)$. However, we continue to assume that $\mu = O(n)$. Inspecting the bias approximation in (3.3) in this case the leading term

is $\frac{\sigma_{uv}}{\sigma_v^2} \frac{k}{(\mu+k)} = O(1)$ and 2SLS is inconsistent. The Hahn and Hausman (2002) bias approximation measures the inconsistency, but note that the remainder term in this case is $O(n^{-1})$ instead of $o(n^{-1})$.

Another possibility is the case of weak instruments (Staiger and Stock, 1997). More in particular, assuming that k is finite, $\pi = c/\sqrt{n}$ with $0 < c < \infty$ we have that $\mu = O(1)$. Again we have that the 2SLS estimator is inconsistent and we have that $E(c) = O(1)$ again, but now $c - E(c) = O_p(1)$. Also we have that $E(d) = O(1)$ now, but now $d - E(d) = O_p(1)$ too. Hence, we are left with a 2SLS estimation error in which both numerator and denominator are of finite stochastic order of magnitude. Hence, an expansion of the type above breaks down.

However, we can combine both cases and consider many weak instruments as introduced by Chao and Swanson (2005). Assuming that $k \rightarrow \infty$ as $n \rightarrow \infty$ with $\mu = O(k)$, which is basically the same result as in the many instruments case. Hence, again the Hahn and Hausman (2002) approximation can be considered as the inconsistency of the IV estimator.

5. Monte Carlo results

The bias approximation in (3.3) is a higher-order approximation irrespective of the number and strength of instruments. Although it breaks down in case of a finite number of weak instruments, it might still be informative in this case. In general it is expected to be more accurate compared with existing bias approximations (2.5) and (2.7).

We generate data on u_i , v_i and z_i as

$$\begin{aligned} \begin{pmatrix} u_i \\ v_i \end{pmatrix} &\sim IIN \left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right) \\ z_i &\sim IIN(0, I_k), \end{aligned}$$

and for y and x according to (2.1) and (2.2). We choose $\rho = 0.5$ and $n = 200$. We evaluate the 2SLS bias approximations for different values of k and μ . We vary the population first stage F statistic μ/k from $\mu/k = 0$ to $\mu/k = 10$ with $k = \{5, 15, 25\}$. We choose the reduced form parameters according to (2.6). With the particular choice for the DGP of the instruments this leads to $\pi_j = \sqrt{\frac{\mu}{kn}}$, $j = 1, \dots, k$.

Figures 1, 2 and 3 shows simulation results for $k = 5$, $k = 15$ and $k = 25$. We plot actual bias (labeled bias) and the 3 different approximations (2.5), (2.7) and (3.3) (labeled Nagar, HH and BW) against the population first stage F statistic μ/k . It is seen that in case of weak instruments the Nagar approximation is inaccurate as documented previously by Cruz and Moreira (2005). For modest k the Hahn-Hausman approximation is inaccurate. The approximation (3.3) works well in all cases.

6. Concluding remarks

In this study the accuracy of various analytical approximations for the finite sample bias of the 2SLS coefficient estimator in the linear instrumental variables model have been analyzed. Using asymptotic expansion techniques an alternative bias approximation has been developed encompassing existing bias approximations. The developed bias approximation is valid in case of both many and many weak instruments.

Through Monte Carlo experiments its accuracy in finite samples is compared with existing bias approximations. The simulation results show that the proposed bias approximation is very accurate for a wide range of parametrizations. In case of many instruments or many weak instruments the proposed bias approximation is numerically very close to an estimate of the inconsistency of the 2SLS estimator. However, in case of only a moderate number of weak instruments invoking higher-order terms yields some significant improvements.

References

- Andrews, D., Stock J.H., 2006. Inference with weak instruments, Chapter 6 in R. Blundell, W.K. Newey and T. Persson, eds., *Advances in Economics and Econometrics, Theory and Applications*, 9th Congress of the Econometric Society, Vol. 3. Cambridge, UK: Cambridge University Press.
- Bekker, P. A., 1994. Alternative approximations to the distributions of instrumental variable estimators. *Econometrica* 62, 657-681.
- Bound, J., Jaeger, D.A., Baker, R.M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association* 90, 443-450.
- Buse, A., 1992. The bias of instrumental variable estimators. *Econometrica* 60, 173-180.
- Chao, J.C., Swanson, N.R., 2005. Consistent estimation with a large number of weak instruments. *Econometrica* 73, 1673-1692.
- Cruz, L.M., Moreira, M.J., 2005. On the validity of econometric techniques with

- weak instruments. *Journal of Human Resources* 40, 393-410.
- Hahn, J., Hausman, J., 2002. Note on bias in estimators for simultaneous equation models. *Economics Letters* 75, 237-241.
- Hahn, J., Kuersteiner, G., 2002. Discontinuities of weak instrument limiting distributions. *Economics Letters* 75, 325-331.
- Nagar, A.L., 1959. The bias and moment matrix of the general k -class estimators of the parameters in simultaneous equations. *Econometrica* 27, 575-595.
- Rothenberg, T.J., 1984. Approximating the distributions of econometric estimators and test statistics. In Z. Griliches and M.D. Intriligator (Eds.), *Handbook of Econometrics, Volume 2*, 881-935. Amsterdam: North Holland.
- Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. *Econometrica* 65, 557-586.
- Stock, J.H., Wright, J.H., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics* 20, 518-529.
- Stock, J.H., Yogo, M., 2005. Testing for weak instruments in linear IV regression. In D.W.K. Andrews and J.H. Stock (Eds.), *Identification and Inference for Econometric Models, Essays in Honor of Thomas Rothenberg*, 80-108. New York: Cambridge University Press.

Figure 1. Bias of 2SLS Estimator and Approximations, $n = 200, \rho = 0.5, k = 5$

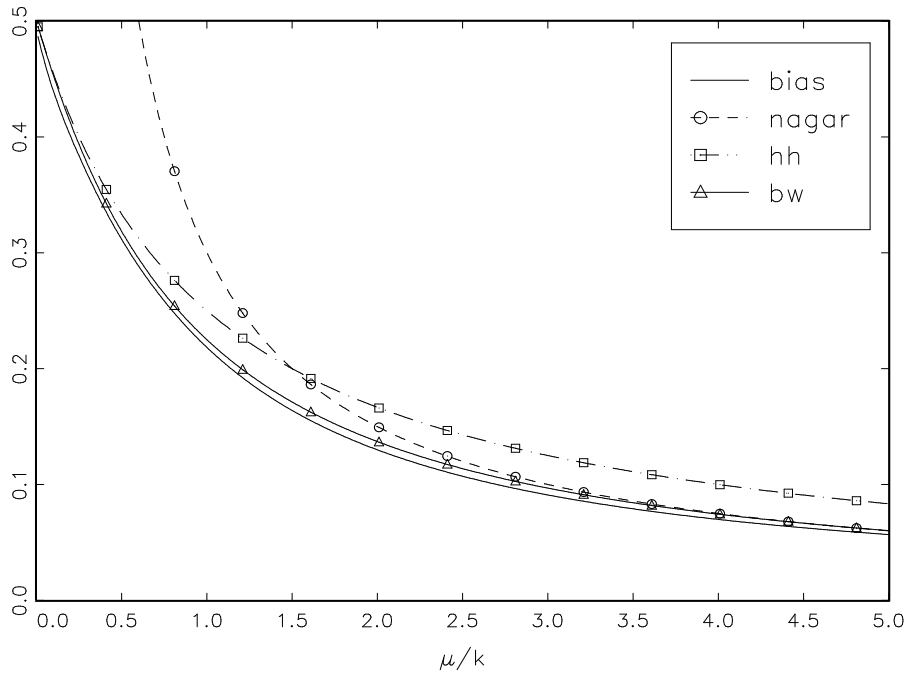


Figure 2. Bias of 2SLS Estimator and Approximations, $n = 200, \rho = 0.5, k = 15$

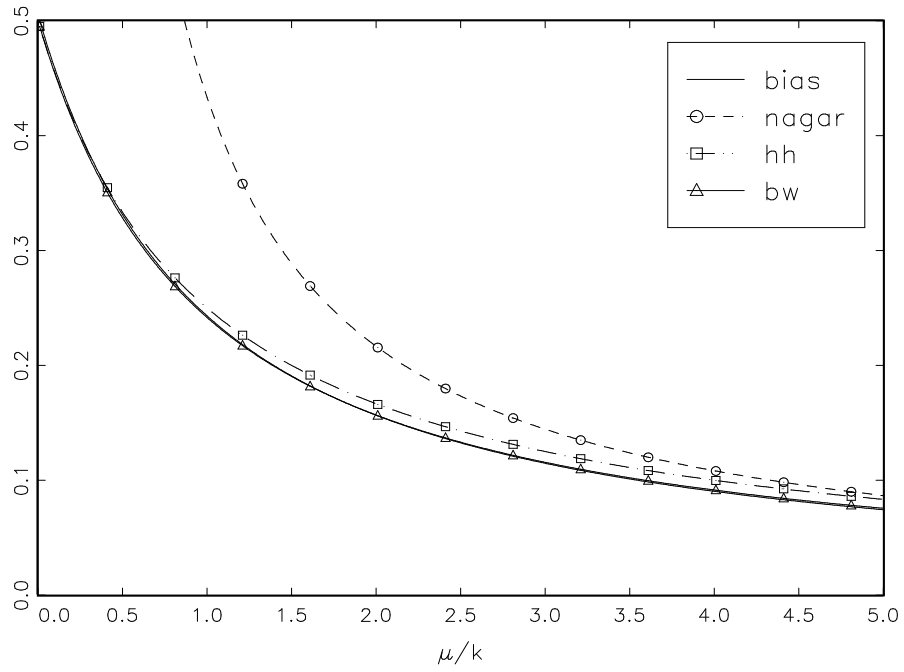


Figure 3. Bias of 2SLS Estimator and Approximations, $n = 200, \rho = 0.5, k = 25$

